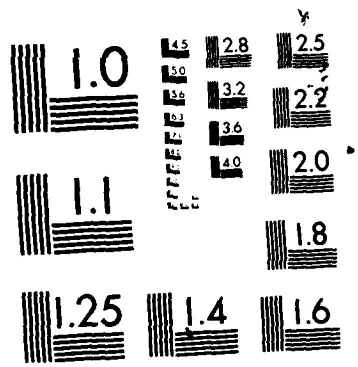
AD-A172 773 B(ASIC)-SPLINE BASICS(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER C DE BOOR AUG 86 MRC-TSR-2952 DAAG29-80-C-0041 1/1

UNCLASSIFIED F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MRC Technical Summary Report #2952

B(asic)-SPLINE BASICS

Carl de Boor

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

August 1986

(Received June 16, 1986)

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

B(asic)-SPLINE BASICS

Carl de Boor

Technical Summary Report #2952
August 1986

ABSTRACT

An expository and illustrated treatment of the basic B-spline theory as derived from the recurrence relations.

AMS (MOS) Subject Classifications: 41A15, 65D10

Key Words: splines, B-splines

Work Unit Number 3 - Numerical Analysis and Scientific Computing

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

SIGNIFICANCE AND EXPLANATION

This report contains the lecture notes for the first of four lectures which comprise the course entitled "The extension of B-spline curve algorithms to surfaces" given at SIGGRAPH'86. It is an elaboration and extension of the MRC report #2896 by de Boor and Höllig, in which the basic B-spline theory is developed from the recurrence relation rather than the original definition in terms of divided differences of the truncated power. This avoids what, to the people in CAGD, amounts to a detour through the theory of divided differences. Somewhat surprisingly, the resulting development is no longer than the standard one, and in some respects seems even more direct. It does bring to the fore the dual functionals and stresses the point that B-splines are best treated in terms of their linear span.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

CONTENTS

B(asic)-SPLINE BASICS

Carl de Boor

0. Introduction

These lecture notes review those basic facts about (univariate) B-splines which are of interest in CAGD. The intent is to give a self-contained and complete development of the material in as simple and direct a way as possible. For this reason, the B-splines are defined via the recurrence relations, thus avoiding the discussion of divided differences which the traditional definition of a B-spline as a divided difference of a truncated power function requires. As this lecture is intended to show, this does not force more elaborate derivations than are available to those who feel at ease with divided differences. It does force a change in the order in which facts are derived and brings more prominence to such things as Marsden's Identity or the Dual Functionals than they currently have in CAGD.

In addition, it highlights the following point: The consideration of a single B-spline is not very fruitful when proving facts about B-splines, even if these facts (such as the smoothness of a B-spline) can be stated in terms of just one B-spline. Rather, simple arguments and real understanding of B-splines are available only if one is willing to consider *all* the B-splines of a given order for a given knot sequence. Thus it focuses attention on **splines**, i.e., on the linear combination of B-splines.

The lecture deals with splines for an arbitrary knot sequence and does rarely become more specific. In particular, the B(ernstein-Bézier)-net for a piecewise polynomial, though a (very) special case of a representation by B-splines, gets much less attention than it deserves, given its immense usefulness in CAGD (and spline theory). But the third lecture takes up this topic.

The lecture deals only with spline **functions**. There is an immediate extension to spline **curves**: Allow the coefficients, be they B-spline coefficients or coefficients in some polynomial form, to be points in $\mathbb{R}^2$ or $\mathbb{R}^3$. But this misses the much richer structure for spline curves available because even discontinuous parametrizations may describe a smooth curve. This topic of **geometric continuity** is discussed in detail in the fourth lecture.

The lecture notes are solidly based on [BH86] which covers more or less the same material, in a less elaborate way and without any figures, in just seven pages.

The relevant literature on (univariate) B-splines up to about 1975 is summarized in [B76] which also contains hints of the most exciting developments concerning B-splines since then: knot insertion and the multivariate B-splines. These are covered in the second lecture, but knot insertion is already put to good use in the last part of this lecture. The two books on splines, [B78] and [Schu81], which have appeared since 1975, cover B-splines in the traditional way. As presentations of splines from the CAGD point of view, the survey article [BFK84] and the "Killer B's" [BBB85,86] are particularly recommended.

Sponsored by the United States Army under Contract No. DAAG29-80-0041.

1. B-splines defined

We start with a partition or **knot sequence**. i.e., a **nondecreasing** sequence $\mathbf{t} := (t_i)$. The **B-splines of order 1** for this knot sequence are the characteristic functions of this partition. i.e., the functions

$$B_{i1}(t) := \mathrm{X}_i(t) := \begin{cases} 1, & \text{if } t_i \le t < t_{i+1} \\ 0, & \text{otherwise.} \end{cases} \tag{1.1}$$

Note that all these functions have been chosen here to be **right-continuous**. Other choices could have been made with equal justification. The only constraint is that these B-splines should form a **partition of unity**, i.e.,

$$\sum_i B_{i1}(t) = 1, \quad \text{for all } t. \tag{1.2}$$

In particular,

$$t_i = t_{i+1} \quad \Longrightarrow \quad B_{i1} = \mathrm{X}_i = 0. \tag{1.3}$$

From these first-order B-splines, we obtain **higher-order** B-splines by **recurrrence**:

$$B_{ik} := \omega_{ik} B_{i,k-1} + (1 - \omega_{i+1,k}) B_{i+1,k-1} \tag{1.4a}$$

with

$$\omega_{ik}(t) := \begin{cases} \frac{t - t_i}{t_{i+k-1} - t_i}, & \text{if } t_i \ne t_{i+k-1} \\ 0, & \text{otherwise.} \end{cases} \tag{1.4b}$$

Thus, the **second-order** B-spline is given by

$$B_{i2} = \omega_{i2} \mathrm{X}_i + (1 - \omega_{i+1,2}) \mathrm{X}_{i+1}, \tag{1.5}$$

and so consists, in general, of two nontrivial linear pieces which join continuously to form a piecewise linear function which vanishes outside the interval $[t_i, t_{i+2}[$. For this reason, some call B_{i2} a **linear** B-spline. If, e.g., $t_i = t_{i+1}$ (hence $\mathrm{X}_i = 0$), but still $t_{i+1} < t_{i+2}$, then B_{i2} consists of just one nontrivial piece and fails to be continuous at the **double knot** $t_i = t_{i+1}$, as is shown in Fig. 1.1.

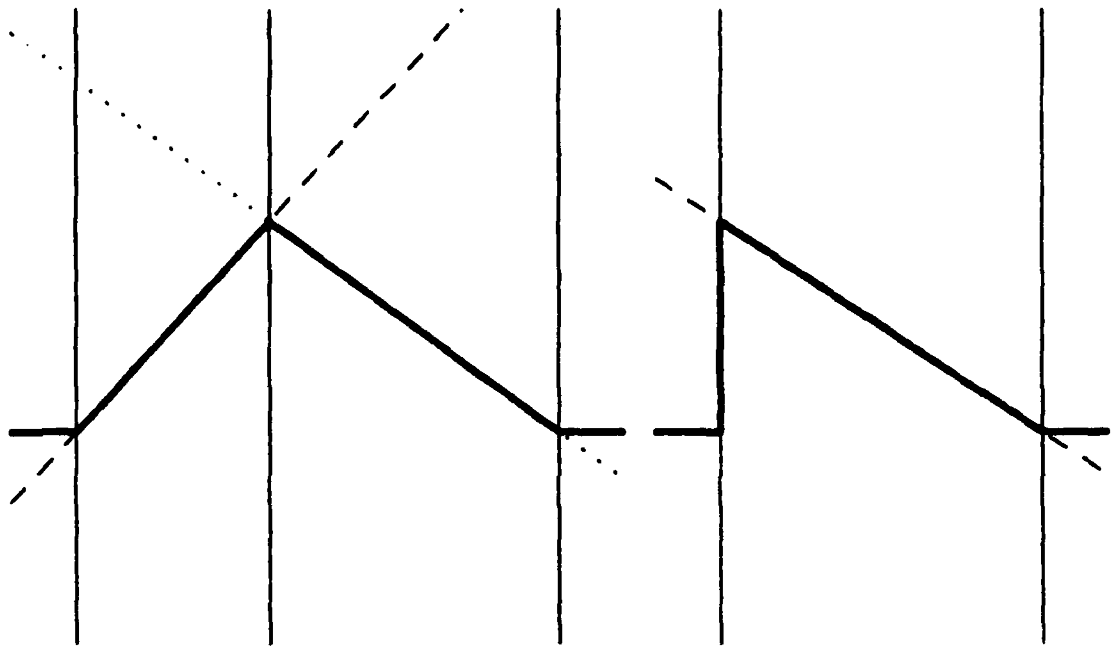


Figure 1.1 Linear B-spline with (a) simple, (b) double knots

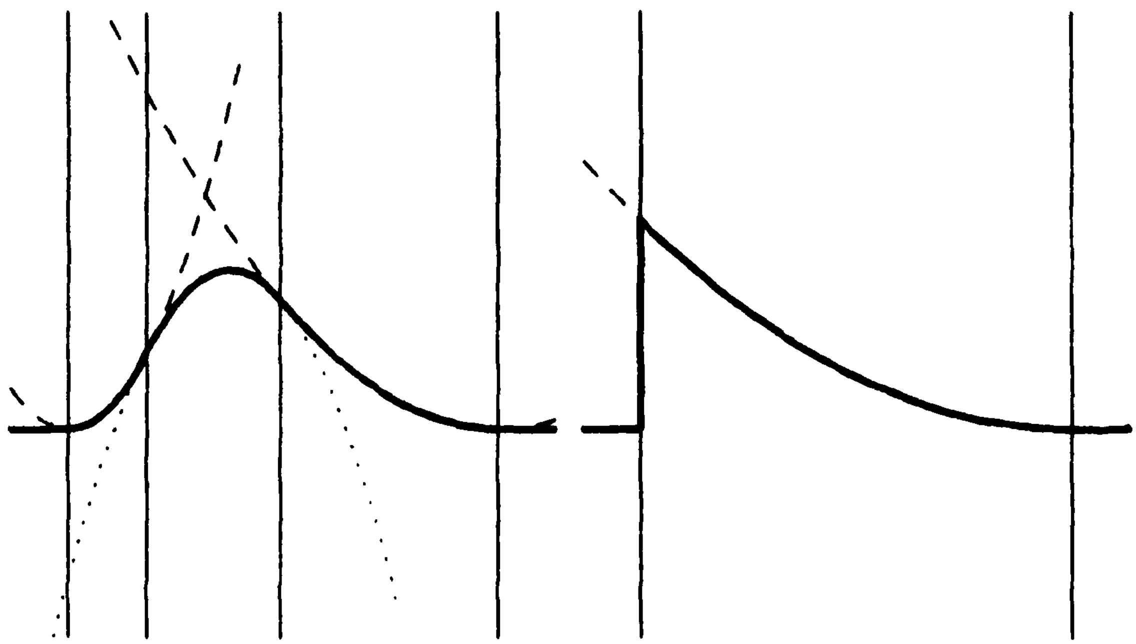


Figure 1.2 Quadratic B-spline with (a) simple, (b) triple knots

The **third-order** B-spline is given by

$$\begin{aligned} B_{i3} &= \omega_{i3}B_{i2} + (1 - \omega_{i+1,3})B_{i+1,2} \\ &= \omega_{i3}\omega_{i2}\chi_i + (\omega_{i3}(1 - \omega_{i+1,2}) + (1 - \omega_{i+1,3})\omega_{i+1,2})\chi_{i+1} \\ &\quad + (1 - \omega_{i+1,3})(1 - \omega_{i+2,2})\chi_{i+2} \end{aligned} \tag{1.6}$$

This shows that, in general, B_{i3} consists of 3 (nontrivial) quadratic pieces, and, to judge from the Fig. 1.2, these seem to join smoothly at the knots to form a C^1 piecewise quadratic function which vanishes outside the interval $[t_i, t_{i+3}]$. Coincidences among the knots $t_i, \ldots, t_{i+3}$ would change this. If, e.g., $t_i = t_{i+1} = t_{i+2}$ (hence $\chi_i = \chi_{i+1} = 0$), then B_{i3} consists of just one nontrivial piece, fails to be even continuous at the **triple knot** t_i, but is still C^1 at the simple knot t_{i+3}, as is shown in Fig. 1.2.

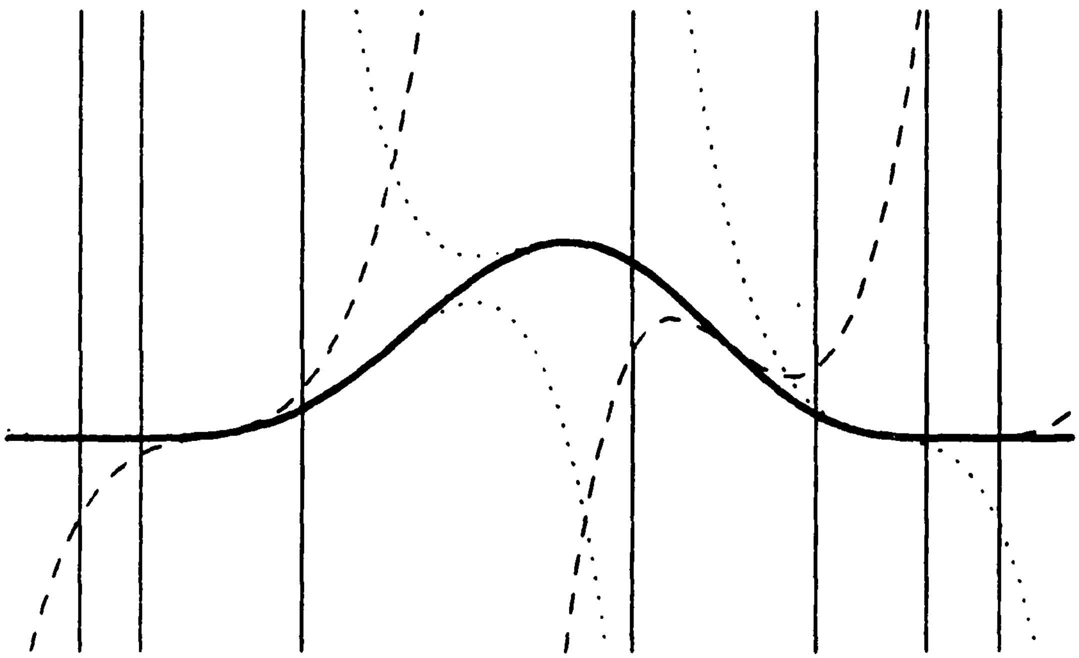


Figure 1.3 A k-order B-spline

After $k - 1$ steps of the recurrence, we obtain B_{ik} in the form

$$B_{ik} = \sum_{j=i}^{i+k-1} b_{jk}\chi_j. \tag{1.7}$$

with each b_{jk} a polynomial of degree $< k$ since it is the sum of products of $k - 1$ linear polynomials.

From this, we infer that B_{ik} is a piecewise polynomial of degree $< k$ which vanishes outside the interval $[t_i, t_{i+k}]$ and has possible **breakpoints** $t_i, \cdots, t_{i+k}$. In particular, B_{ik}

is just the zero function in case $t_i = t_{i+k}$. Also, by induction, B_{ik} is positive on the open interval $]t_i, t_{i+k}[$, since both ω_{ik} and $1 - \omega_{i+1,k}$ are positive there.

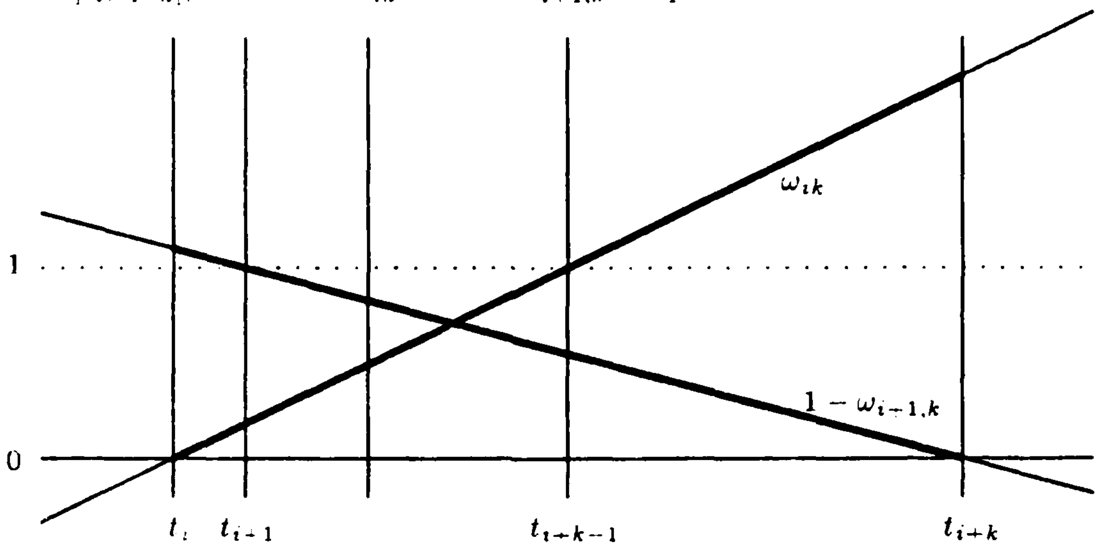


Figure 1.4 The two weight functions in (1.4a) are positive on $]t_i, t_{i+k}[=$ supp B_{ik}.

Further, we see that B_{ik} is completely determined by the $k+1$ knots $t_i, \ldots, t_{i+k}$. For this reason, the notation

$$B(\cdot|t_i, \ldots, t_{i+k}) \;:=\; B_{ik} \tag{1.8}$$

is sometimes used. Other notations in use include

$$N_{ik} := B_{ik} \text{ and } M_{ik} := \big(k/(t_{i+k} - t_i)\big)B_{ik}. \tag{1.9}$$

The many other properties of B-splines are derived most easily by considering not just one B-spline but the linear span of all B-splines of a given order k for a given knot sequence $\mathbf{t}$. This brings us to splines.

2. Splines defined

A **spline of order k with knot sequence $\mathbf{t}$** is, by definition, a linear combination of the B-splines B_{ik} associated with that knot sequence. We denote by

$$S_{k,\mathbf{t}} := \{\sum_i B_{ik}a_i \;:\; a_i \in \mathbb{R}\} \tag{2.1}$$

the collection of all such splines.

We have left open so far the precise nature of the knot sequence $\mathbf{t}$, other than to specify that it be a nondecreasing real sequence. In any practical situation, $\mathbf{t}$ is necessarily a *finite* sequence. But, since on any nontrivial interval $[t_j, t_{j+1}[$ at most k of the B_{ik} are nonzero, viz. $B_{j-k+1,k}, \ldots, B_{jk}$, it doesn't really matter whether $\mathbf{t}$ is finite, infinite, or even bi-infinite; the sum in (2.1) always makes pointwise sense, since, on any interval $[t_j, t_{j+1}[$, at most k summands are not zero.

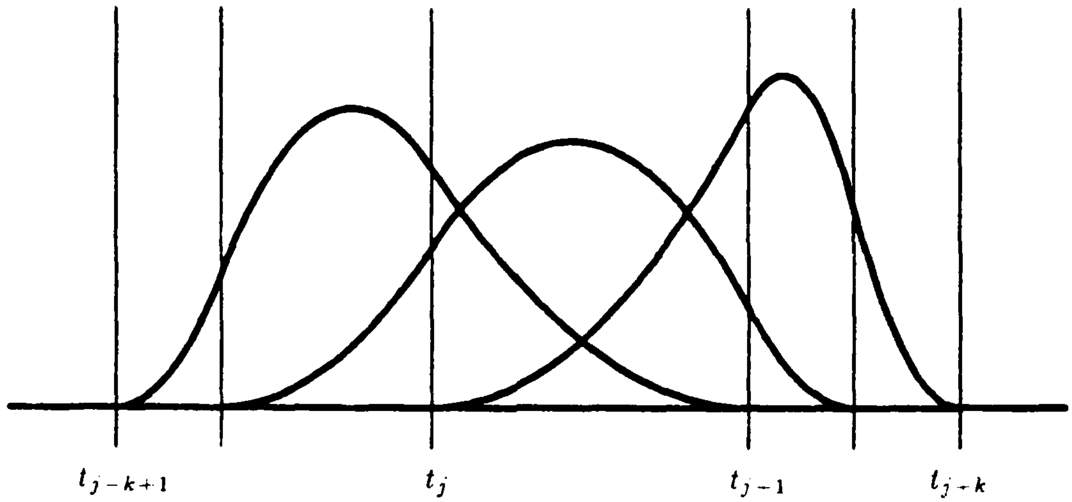


Figure 2.1 The k B-splines whose support contains $[t_j, t_{j+1}[$

We will pay special attention to the following two "extreme" knot sequences, the sequence

$$\mathbb{Z} := (\ldots, -2, -1, 0, 1, 2, \ldots)$$

and the sequence

$$\mathbb{B} := (\ldots, 0, 0, 0, 1, 1, 1, \ldots).$$

A spline associated with the knot sequence $\mathbb{Z}$ is called **cardinal** splines. This term was chosen by Schoenberg [Scho69] because of a connection to Whittaker's Cardinal Series. This is not to be confused with its use in earlier spline literature where it refers to a spline which vanishes at all points in a given sequence except for one at which it takes the value 1. The latter splines, though of great interest in spline interpolation, do not interest us here.

Because of the uniformity of the knot sequence $\mathbf{t} = \mathbb{Z}$, formulae involving cardinal B-splines are often much simpler than corresponding formulae for general B-splines. To begin with, *all cardinal B-splines (of a given order) are translates of one another*. With the natural indexing $t_i := i, \forall i$, for the entries of the uniform knot sequence $\mathbf{t} = \mathbb{Z}$, we have

$$B_{ik} = N_k(\cdot - i), \tag{2.2}$$

with

$$N_k := B_{0k} = B(\cdot|0, \ldots, k). \tag{2.3}$$

The recurrence relations (1.4) simplify as follows:

$$(k-1)N_k(t) = tN_{k-1}(t) + (k-t)N_{k-1}(t-1). \tag{2.4}$$

The knot sequence $\mathbf{t} = \mathbb{B}$ contains just two points, viz., the points 0 and 1, but each with infinite multiplicity. The only nontrivial B-splines for this sequence are those which have both 0 and 1 as knots, i.e., those B_{ik} for which $t_i = 0$ and $t_{i+k} = 1$. There seems to be no natural way to index the entries in the sequence $\mathbb{B}$. Instead, it is customary to index the corresponding B-splines by the multiplicities of their two distinct knots. Precisely,

$$B_{(\mu,\nu)} := B(\cdot|\underbrace{0,\ldots,0}_{\mu+1 \text{ times}},\underbrace{1,\ldots,1}_{\nu+1 \text{ times}}). \tag{2.5}$$

With this, the recurrence relations (1.4) simplify as follows:

$$B_{(\mu,\nu)}(t) = tB_{(\mu,\nu-1)}(t) + (1-t)B_{(\mu-1,\nu)}(t). \tag{2.6}$$

This gives the formula

$$B_{(\mu,\nu)}(t) = \binom{\mu+\nu}{\mu}(1-t)^{\mu}t^{\nu} \quad \text{for } 0 < t < 1 \tag{2.7}$$

for the one nontrivial polynomial piece of $B_{(\mu,\nu)}$, as one verifies by induction. The formula enables us to determine the *smoothness* of the B-splines in this simple case: Since $B_{(\mu,\nu)}$ vanishes identically outside $[0..1]$, it has exactly $\nu - 1$ continuous derivatives at 0 and $\mu - 1$ continuous derivatives at 1. This amounts to ν **smoothness conditions** at 0 and μ smoothness conditions at 1. Since the order of $B_{(\mu,\nu)}$ is $\mu + \nu + 1$, this is a simple illustration of the generally valid formula

$$\#\text{smoothness conditions at knot} + \text{multiplicity of knot} = \text{order}. \tag{2.8}$$

For fixed $\mu + \nu$, the polynomials in (2.7) form the so-called **Bernstein** basis (for polynomials of degree $\le \mu + \nu$) and, correspondingly, the representation

$$p = \sum_{\mu+\nu=h} B_{(\mu,\nu)}a_{(\mu,\nu)} \tag{2.9}$$

is the **Bernstein** form for the polynomial $p \in \pi_h$. In CAGD, it is more customary to refer to (2.9) as the **Bézier** form (for the polynomial p) or as the **Bézier polynomial** or even the **Bernstein-Bézier** polynomial. It may be simpler to use the short term **B-form** instead.

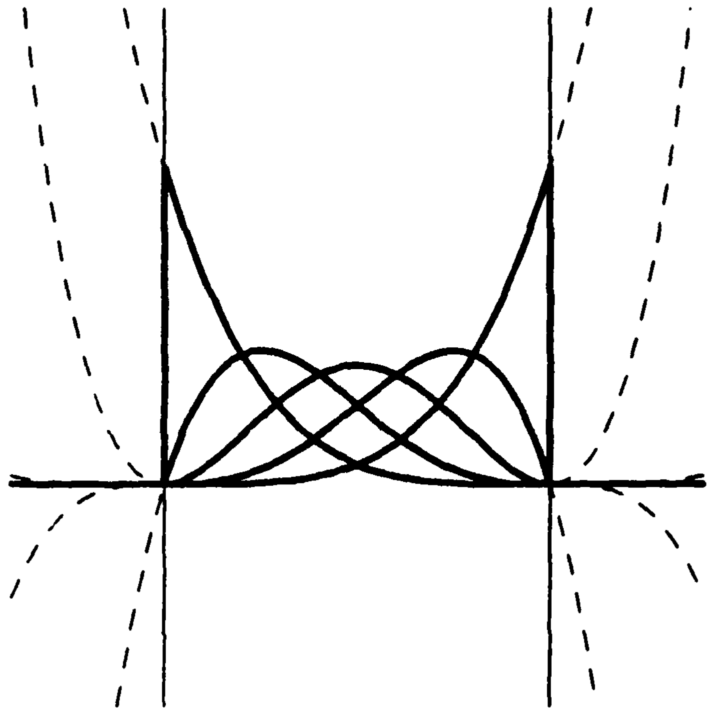


Figure 2.2 Bernstein basis of degree 4

3. A simplifying assumption

In the next sections, we develop the basic B-spline theory by studying the spline space $S_{k,\mathbf{t}}$, i.e., the collection of all functions s of the form

$$s = \sum_i B_{ik} a_i \tag{3.1}$$

for a suitable coefficient vector $a = (a_i)$.

In practice, the knot sequence $\mathbf{t}$ is always finite, hence so is the sum in (3.1). This often requires one to pay special attention to the limits of that summation. Since I find that distracting, I will assume from now on that *the knot sequence* $\mathbf{t}$ *is bi-infinite*. This can always be achieved simply by continuing the sequence indefinitely in both directions (taking care to maintain its monotonicity) and choosing the additional B-spline coefficients to be zero.

More than that, I will assume that

$$t_{\pm\infty} := \lim_{i \to \pm\infty} t_i = \pm\infty. \tag{3.2}$$

This assumption is convenient since it ensures that every $\tau \in \mathbb{R}$ is in the support of some B-spline.

At times, it will be convenient to assume that

$$t_i < t_{i+k} \quad \forall i \tag{3.3}$$

which can always be achieved by removing from $\mathbf{t}$ its i-th entry while $t_i = t_{i+k}$. This does not change the space $S_{k,\mathbf{t}}$ since the only k-order B-splines removed thereby are zero anyway. In fact, another way to state the condition (3.3) is:

$$B_{ik} \neq 0 \quad \forall i. \tag{3.3'}$$

4. The polynomials in $S_{k,\mathbf{t}}$

We show in this section that $S_{k,\mathbf{t}}$ contains

$$\pi_{<k} := \text{the collection of all polynomials of degree } < k,$$

and give a formula for the B-spline coefficients of $p \in \pi_{<k}$.

We begin with **Marsden's Identity**:

Theorem 4. For any $\tau \in \mathbb{R}$,

$$(\cdot - \tau)^{k-1} = \sum_i B_{ik}\psi_{ik}(\tau), \tag{4.1a}$$

with

$$\psi_{ik}(\tau) := (t_{i+1} - \tau)\cdots(t_{i+k-1} - \tau). \tag{4.1b}$$

Proof We deduce from the recurrence relation (1.4) that, for an arbitrary coefficient sequence a,

$$\sum B_{ik}a_i = \sum B_{i,k-1}\big((1 - \omega_{ik})a_{i-1} + \omega_{ik}a_i\big). \tag{4.2}$$

On the other hand, for the special sequence

$$a_i := \psi_{ik}(\tau) := (t_{i+1} - \tau)\cdots(t_{i+k-1} - \tau)$$

(with $\tau \in \mathbb{R}$), we find for $B_{i,k-1} \neq 0$, i.e., for $t_i < t_{i+k-1}$ that

$$\begin{aligned}(1 - \omega_{ik})a_{i-1} + \omega_{ik}a_i &= \Big((1 - \omega_{ik})(t_i - \tau) + \omega_{ik}\cdot(t_{i+k-1} - \tau)\Big)\psi_{i,k-1}(\tau)\\ &= (\cdot - \tau)\psi_{i,k-1}(\tau)\end{aligned} \tag{4.3}$$

since $(1 - \omega_{ik})f(t_i) + \omega_{ik}f(t_{i+k-1})$ is the straight line which agrees with f at t_i and t_{i+k-1}, hence must equal f if, as in our case, f is linear. This shows that

$$\sum B_{ik}\psi_{ik}(\tau) = (\cdot - \tau)\sum B_{i,k-1}\psi_{i,k-1}(\tau).$$

hence, by induction, that

$$\sum B_{ik}\psi_{ik}(\tau) = (\cdot - \tau)^{k-1} \sum B_{i1}\psi_{i1}(\tau) = (\cdot - \tau)^{k-1},$$

since $\psi_{i1}(\tau) = 1$ and $\sum_i B_{i1} = 1$ (see (1.2)). |||

Remark There may be some doubt as to why ψ_{i1} should be identically equal to 1. From the definition (4.1b), it would appear that ψ_{i1} is the product of *no* factors, hence, by a standard agreement concerning the empty product, equal to 1. This is the definition appropriate for use in induction arguments. Indeed, if you consider the coefficients in (4.2) for $k = 2$ directly, you get

$$\begin{aligned}(1 - \omega_{i2})\psi_{i-1,2}(\tau) + \omega_{i2}\psi_{i2}(\tau) &= (1 - \omega_{i2}) \cdot (t_i - \tau) + \omega_{i2} \cdot (t_{i+1} - \tau) \\ &= (\cdot - \tau),\end{aligned}$$

which agrees with (4.3) for this case if we set $\psi_{i1}(\tau) = 1$.

Since τ in (4.1) is arbitrary, it follows that $S_{k,\mathbf{t}}$ contains all polynomials of degree $< k$. More than that, we can even give an explicit expression for the required coefficients, as follows.

By dividing (4.1a) by $(k-1)!$ and then differentiating it with respect to τ, we obtain the identities

$$\frac{(\cdot - \tau)^{k-\nu}}{(k-\nu)!} = \sum_i B_{ik} \frac{(-D)^{\nu-1}\psi_{ik}(\tau)}{(k-1)!}, \quad \nu > 0, \tag{4.4}$$

with Df the derivative of the function f. On using this identity in the Taylor formula

$$p = \sum_{\nu=1}^{k} \frac{(\cdot - \tau)^{k-\nu}}{(k-\nu)!} D^{k-\nu} p(\tau),$$

valid for any $p \in \pi_{<k}$, we conclude that any such polynomial can be written in the form

$$p = \sum_i B_{ik}\lambda_{ik} p. \tag{4.5a}$$

with λ_{ik} given by the rule

$$\lambda_{ik} f := \sum_{\nu=1}^{k} \frac{(-D)^{\nu-1}\psi_{ik}(\tau)}{(k-1)!} D^{k-\nu} f(\tau). \tag{4.5b}$$

Here are two special cases of particular interest. For $p = 1$, we get

$$1 = \sum_i B_{ik} \tag{4.6}$$

since $D^{k-1}\psi_{ik} = (-1)^{k-1}(k-1)!$, and this shows that the B_{ik} form a **partition of unity**. Further, since $D^{k-2}\psi_{ik}$ is a linear polynomial which vanishes at

$$t_i^* := (t_{i+1} + \cdots + t_{i+k-1})/(k-1), \tag{4.7}$$

we get the important identity

$$p = \sum_i B_{ik} p(t_i^*) \quad \forall p \in \pi_1. \tag{4.8}$$

Remark In the *cardinal* case,

$$\psi_{ik}(\tau)/(k-1)! = \binom{i - \tau + k - 1}{k-1}, \tag{4.1b)$_{\mathbb{Z}}$$

while in the *Bernstein-Bézier* case,

$$\psi_{(\mu,\nu)}(\tau) = (-\tau)^{\mu}(1-\tau)^{\nu} = (-)^{\mu} B_{(\nu,\mu)} / \binom{\mu+\nu}{\mu}. \qquad (4.1b)_{\mathbb{B}}$$

5. The pp functions contained in $S_{k,\mathbf{t}}$

In this section, we show that the spline space $S_{k,\mathbf{t}}$ coincides with a certain space of **pp(:= piecewise polynomial)** functions.

Each $s \in S_{k,\mathbf{t}}$ is pp of degree $< k$ with breakpoint sequence $\mathbf{t}$ since each B_{ik} is pp of degree $< k$ and has breakpoints $t_i, \ldots, t_{i+k}$. In symbols,

$$S_{k,\mathbf{t}} \subseteq \pi_{<k,\mathbf{t}}. \tag{5.1}$$

But $S_{k,\mathbf{t}}$ is usually a proper subspace of $\pi_{<k,\mathbf{t}}$ since, depending on the **knot multiplicities**

$$\#t_i := \#\{t_j : t_i = t_j\}, \tag{5.2}$$

the splines in $S_{k,\mathbf{t}}$ are more or less smooth, while the typical element of $\pi_{<k,\mathbf{t}}$ has jump discontinuities at every t_i.

For the precise description, given in Theorem 5 below, of the **smoothness conditions** satisfied by the elements of $S_{k,\mathbf{t}}$, we make use of Marsden's Identity, (4.1), since it provides us with the B-spline coefficients of various *pp* functions in $S_{k,\mathbf{t}}$, as follows.

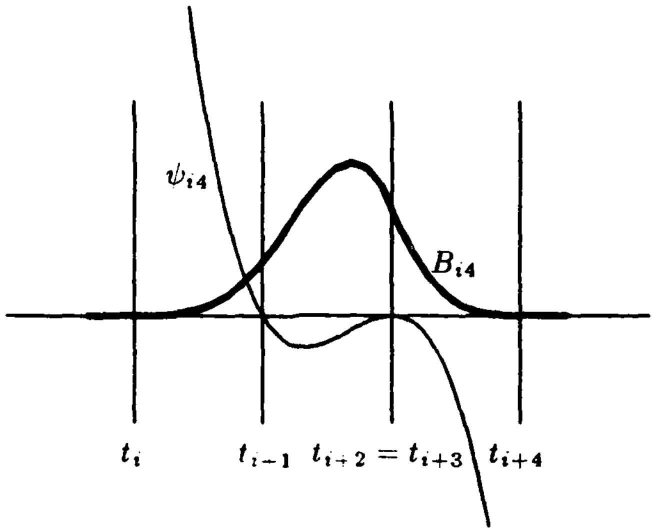


Figure 5.1 B_{i4} and ψ_{i4}; note the double knot

Since $B_{ik}(t_j) \neq 0$ implies $\psi_{ik}(t_j) = 0$ (see Fig. 5.1), the choice $\tau = t_j$ in (4.1) leaves only terms with support either entirely to the left or else entirely to the right of t_j; see Fig. 5.2. This implies that

$$(\cdot - t_j)_+^{k-1} = \sum_{i \geq j} B_{ik}\psi_{ik}(t_j) \tag{5.3}$$

with

$$\alpha_+ := \max\{\alpha, 0\} \tag{5.4}$$

the positive part of the number α. More than that, since $B_{ik}(t_j) \neq 0$ implies $D^{\nu-1}\psi_{ik}(t_j) = 0$ in case $\nu \leq \#t_j$, the same observation applied to (4.4) shows that

$$(\cdot - t_j)_+^{k-\nu} \in S_{k,\mathbf{t}} \quad \text{for} \quad 1 \leq \nu \leq \#t_j. \tag{5.5}$$

Theorem 5. The space $S_{k,\mathbf{t}}$ coincides with the space $\tilde{S}$ of all piecewise polynomials of degree $< k$ with breakpoints t_i which are $k - 1 - \#t_i$ times continuously differentiable at t_i.

Proof Assume without loss of generality (see Sec. 3) that

$$t_i < t_{i+k} \quad \forall i.$$

It is sufficient to prove that, for any finite interval $I := [a, b]$, the restriction $\tilde{S}_{|I}$ of the space $\tilde{S}$ to the interval I coincides with the restriction of $S_{k,\mathbf{t}}$ to that interval. The latter space is spanned by all the B-splines having some support in I, i.e., all B_{ik} with $(t_i, t_{i+k}) \cap I \neq \emptyset$. The space $\tilde{S}_{|I}$ has a basis consisting of the functions

$$(\cdot - a)^{k-\nu},\ \nu = 1, \cdots, k;\ (\cdot - t_i)_+^{k-\nu},\ \nu = 1, \cdots, \#t_i,\ \text{for}\ a < t_i < b. \tag{5.6}$$

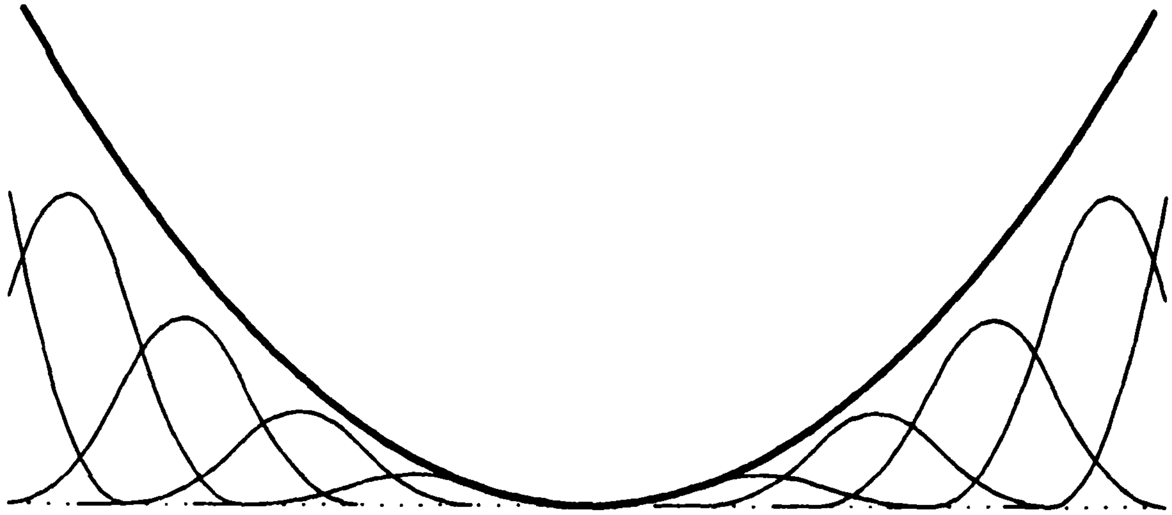


Figure 5.2 The summands $B_{i3}\psi_{i3}(t_j)$ and their sum, $(\cdot - t_j)^2$

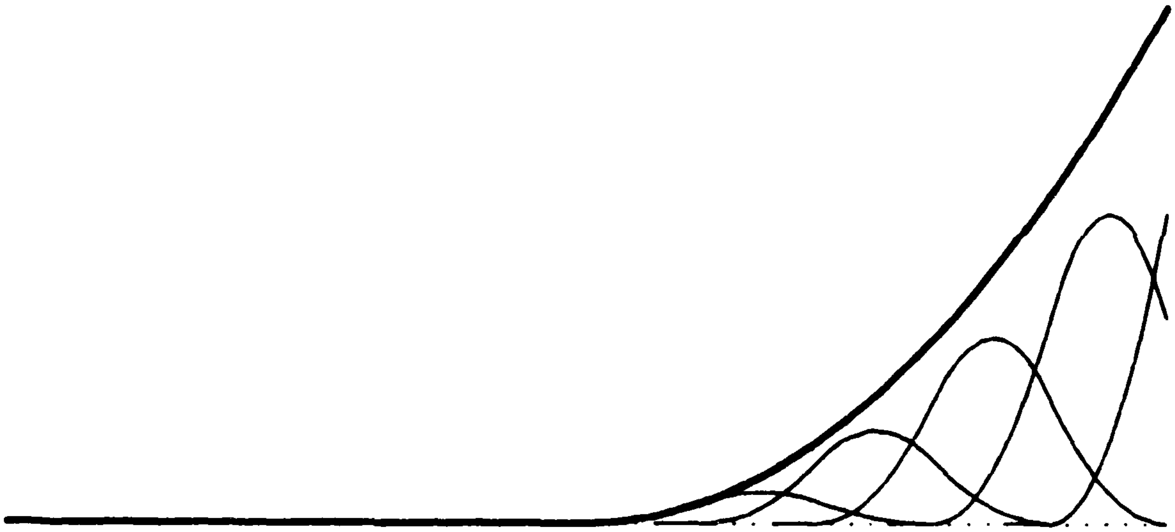


Figure 5.3 The summands $B_{i3}\psi_{i3}(t_j), i \geq j$, and their sum, $(\cdot - t_j)_+^2$

This follows from the observation that a piecewise polynomial function f with a breakpoint at t_i which is $k - 1 - \#t_i$ times continuously differentiable there can be written uniquely as

$$f = p + \sum_{\nu=1}^{\#t_i} (\cdot - t_i)_+^{k-\nu} c_\nu,$$

with p a suitable polynomial of degree $< k$ and suitable coefficients c_ν. Since each of the functions in (5.6) lies in $S_{k,\mathbf{t}}$, by (5.3) and (5.5), we conclude that

$$\tilde{S}_{|I} \subseteq (S_{k,\mathbf{t}})_{|I}. \tag{5.7}$$

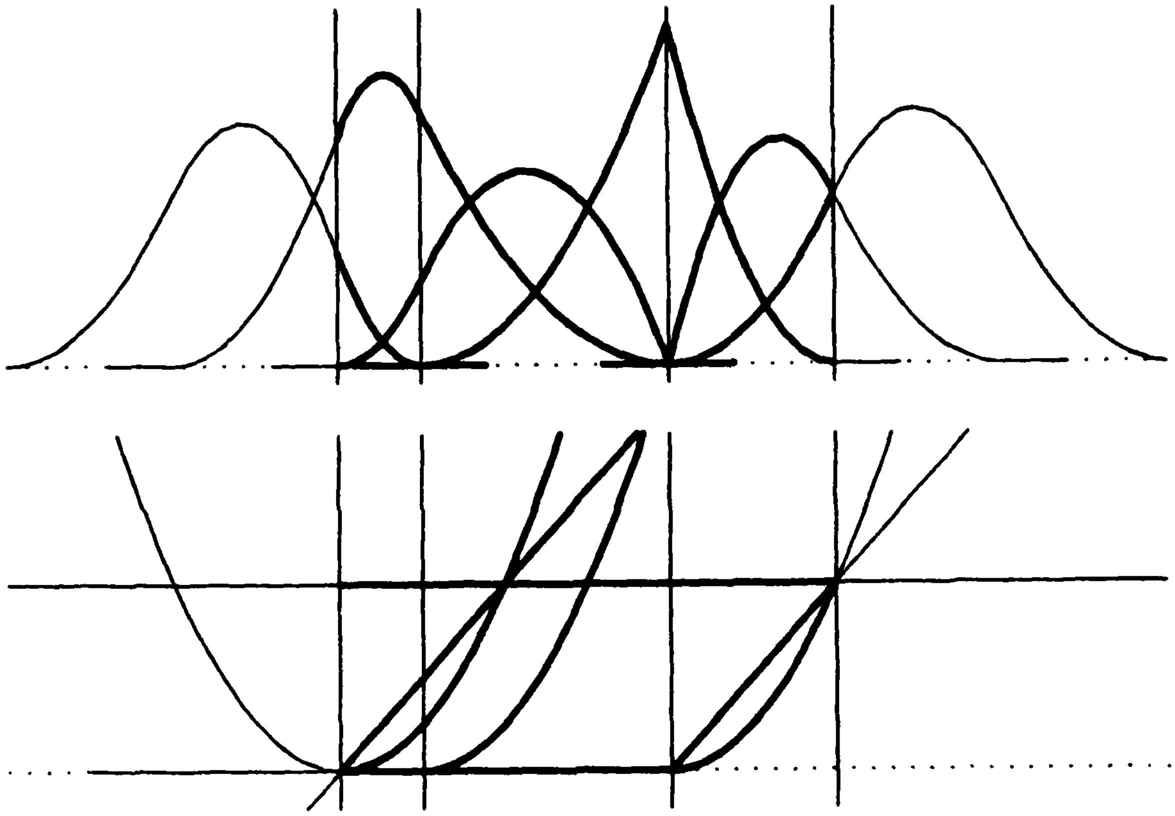


Figure 5.4 (a) The six quadratic B-splines for the case of one simple and one double interior knot; and
(b) the corresponding truncated power basis.

On the other hand, the dimension of $\tilde{S}_{|I}$, i.e., the number of functions in (5.6), equals the number of B-splines with some support in I (since it equals $k + \sum_{a<t_i<b} \#t_i$), hence is an upper bound on the dimension of $(S_{k,\mathbf{t}})_{|I}$. This implies that equality must hold in (5.7), which is what we set out to prove. |||

Remark The argument from Linear Algebra used here is the following: Suppose that we know a basis, $(f_1, f_2, \ldots, f_n)$ say, for the linear subspace F, and that we further know a sequence $(g_1, g_2, \ldots, g_m)$ whose span, G say, contains each of the f_i. Then, of course, $F \subseteq G$ and so

$$n = \dim F \leq \dim G \leq m.$$

If we now know, in addition, that $n = m$, then necessarily $F = G$. Moreover, then necesarily $\dim G = m$, i.e., the sequence $(g_1, g_2, \ldots, g_m)$ must be linearly independent (since it then is *minimally spanning* for G). In our particular situation, this last observation implies that the set of B-splines having some support in I must be linearly independent over I. We pick up on this in the next section.

6. 'B' stands for 'BASIC'

In this section, we discuss the **basis property** of the B-splines, as a consequence of Theorem 5 and its proof.

From the Remark following Theorem 5, we obtain the following *sharpening* of Theorem 5.

Theorem 6. Let $I := [a, b]$ be a finite interval. Then the restrictions

$$\{B_{ik|I} : B_{ik|I} \neq 0\} \tag{6.1}$$

of those B-splines which have some support on I form a basis for the space of pp functions of degree $< k$ on I with breakpoints $\{t_i : a < t_i < b\}$ and which are $k - 1 - \#t_i$ continuously differentiable at each of their breakpoints t_i.

We conclude that the number of smoothness conditions at a knot t_i guaranteed to be satisfied by every spline in $S_{k,\mathbf{t}}$ equals $k - \#t_i$. This proves the formula

$$\#\text{smoothness conditions at knot} \;+\; \text{multiplicity of knot} \;=\; \text{order} \tag{2.8}$$

cited earlier (in connection with the Bernstein-Bézier form).

It is worthwhile to think about this the other way around. Suppose we start off with a partition

$$a =: \xi_1 < \xi_2 < \cdots < \xi_\ell < \xi_{\ell+1} := b$$

of the interval $I := [a, b]$ and wish to consider the space

$$\pi^{\nu}_{<k,\xi}$$

of all pp functions of degree $< k$ on I with breakpoints ξ_i which satisfy ν_i smoothness conditions at ξ_i, i.e., are $\nu_i - 1$ times continuously differentiable at $\xi_i, \forall i$. Then a B-spline basis for this space is provided by (6.1), with the knot sequence $\mathbf{t}$ constructed from the breakpoint sequence ξ in the following way: To the sequence

$$(\underbrace{\xi_2, \ldots, \xi_2}_{\nu_2 \text{ times}}, \underbrace{\xi_3, \ldots, \xi_3}_{\nu_3 \text{ times}}, \ldots, \underbrace{\xi_\ell, \ldots, \xi_\ell}_{\nu_\ell \text{ times}}). \tag{6.2}$$

adjoin at the beginning k points $\leq a$ and at the end k points $\geq b$. While the knots in (6.2) have to be exactly as shown to achieve the specified smoothness at the specified breakpoints, the $2k$ additional knots are quite arbitrary. They are often chosen to equal a resp. b, and this has certain advantages (among other things that of simplicity). With such a choice, it is necessary to modify the definition (1.1) so as to include the right endpoint, b, into the support of the rightmost nontrivial B_{i1}. In other words, if n is such that

$$t_n < t_{n+1} = b.$$

then

$$B_{n1}(t) := X_n(t) := \begin{cases} 1, & \text{if } t_i \le t \le b \\ 0, & \text{otherwise.} \end{cases} \tag{6.3}$$

This ensures that, in evaluating a spline or its derivatives at b, we obtain the limit from the left.

The identification of $S_{k,t}$ with a certain space of pp functions allows the following conclusions of importance in calculations to be discussed later.

Corollary 1. If $t_i < t_{i+k-1}$, then the derivative of a spline in $S_{k,t}$ is a spline of degree $< k-1$ with respect to the same knot sequence, i.e., $DS_{k,t} \subsetneq S_{k-1,t}$.

Proof By assumption, $\#t_i < k$, hence the pp functions in $S_{k,t}$ are continuous, therefore differentiable (if we accept a possible jump at t_i in the derivative Ds of $s \in S_{k,t}$ in case $\#t_i = k-1$). Further, such a derivative Ds is pp of degree $< k-1$ and satisfies $k - \#t_i - 1$ smoothness conditions at t_i, hence belongs to $S_{k-1,t}$, by Theorem 5 or 6. |||

Corollary 2. If $\hat{t}$ is a refinement of the knot sequence t, then $S_{k,t} \subset S_{k,\hat{t}}$.

Proof Since $\hat{t}$ is a refinement of t, i.e., contains entries in addition to those of t, the pp functions in $S_{k,t}$ satisfy all the conditions which, by Theorem 5 or 6, characterize the pp functions in $S_{k,\hat{t}}$. (But the converse does not hold, since the pp functions in $S_{k,\hat{t}}$ may have more breakpoints and/or may be less smooth at some breakpoints than the pp functions in $S_{k,t}$.) |||

These corollaries point out that it should be possible, in principle, to compute from the B-spline coefficients of a spline in $S_{k,t}$ the B-spline coefficients of its derivative and its B-spline coefficients with respect to a refined knot sequence. To carry out such calculations, though, we need a means of expressing the B-spline coefficients of a spline in terms of other information, such as its values and derivatives at certain points. If the spline happens to be a polynomial, then such a formula is provided by (4.5). We show in the next section that the same formula works for any spline (provided we are willing to restrict the parameter τ suitably).

7. The dual functionals

In this section, we prove that the formula (4.5) for the B-spline coefficients of a polynomial is valid for an arbitrary spline provided we restrict the parameter τ in the definition

$$\lambda_{ik} : f \mapsto \sum_{\nu=1}^{k} \frac{(-D)^{\nu-1}\psi_{ik}(\tau)}{(k-1)!} D^{k-\nu} f(\tau) \tag{4.5b}$$

to the support of B_{ik}.

For this, we agree, consistent with (1.1b), that all derivatives in (4.5b) are to be taken as limits from the right in case τ coincides with a knot (except, perhaps, when τ is the right endpoint of the interval of interest, see (6.3)).

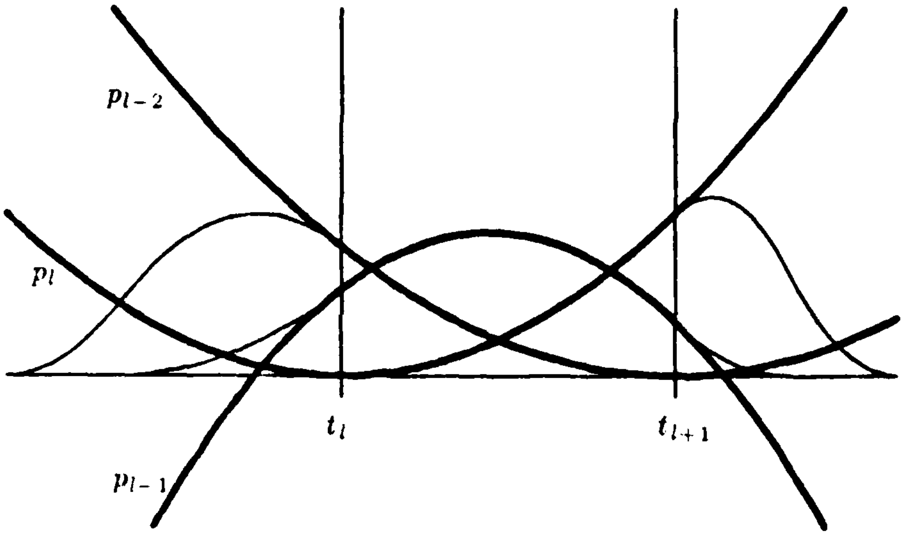


Figure 7.1 The three polynomials, p_{l-2}, p_{l-1}, p_l, which agree with some quadratic B-spline B_{j3} on the knot interval $[t_l, t_{l+1}[$.

Theorem 7. If τ in definition (4.5b) of λ_{ik} is chosen in the interval $[t_i, t_{i+k}]$, then

$$\lambda_{ik}\Big(\sum_j B_{jk}a_j\Big) = a_i. \tag{7.1}$$

Proof We prove that, under the given restriction,

$$\lambda_{ik}B_{jk} = \delta_{ij} := \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{otherwise.} \end{cases} \tag{7.2}$$

Assume that $\tau \in [t_l, t_{l+1}[\subset [t_i, t_{i+k}]$. Then (7.2) requires proof only for $j = l-k+1, \dots, l$ since, for all other j, $i \neq j$ and B_{jk} vanishes identically on $[t_l, t_{l+1}[$, hence also $\lambda_{ik}B_{jk} = 0$. For each of the remaining j's, let p_j be the polynomial which agrees with B_{jk} on $[t_l, t_{l+1}[$. Then

$$\lambda_{ik}B_{jk} = \lambda_{ik}p_j.$$

On the other hand,

$$p_j = \sum_{i=l-k+1}^{l} p_i\, \lambda_{ik}p_j\ . \tag{7.3}$$

since this holds by (4.5a) on $[t_l, t_{l+1}[$. This forces $\lambda_{ik}p_j$, hence $\lambda_{ik}B_{jk}$, to equal δ_{ij} for $i,j = l-k+1,\ldots,l$, since, by Theorem 6 or directly from the fact that (4.5a) holds for every $p \in \pi_{<k}$, the sequence

$$p_{l-k+1},\cdots,p_l \tag{7.4}$$

is linearly independent. |||

Remark The argument used here is that, for a linearly independent sequence $(f_1,\ldots,f_n)$, the only way the equation

$$f_i = \sum_{j=1}^{n} f_j a_{ij}$$

can hold is for a_{ij} to equal 1 for $i = j$ and zero otherwise. Further, the linear independence of the sequence (7.4) follows from the validity of (4.5a) for every $p \in \pi_{<k}$ since that implies that the k-sequence (7.4) is spanning for the k-dimensional space $\pi_{<k}$. It also follows from Theorem 6 with $I = [t_l, t_{l+1}[$.

The two sequences, (B_{ik}) and (λ_{jk}), are said to be **bi-orthonormal** or **dual to each other** because they satisfy (7.2). For this reason, the linear functionals λ_{ik} are at times referred to as the **dual functionals** for the B-splines.

We exploit the simple formula (7.1) for the i-th B-spline coefficient of a spline in subsequent sections, in order to derive algorithms for differentiation and knot insertion and, ultimately, to derive statements about the condition and the shape-preserving property of B-splines.

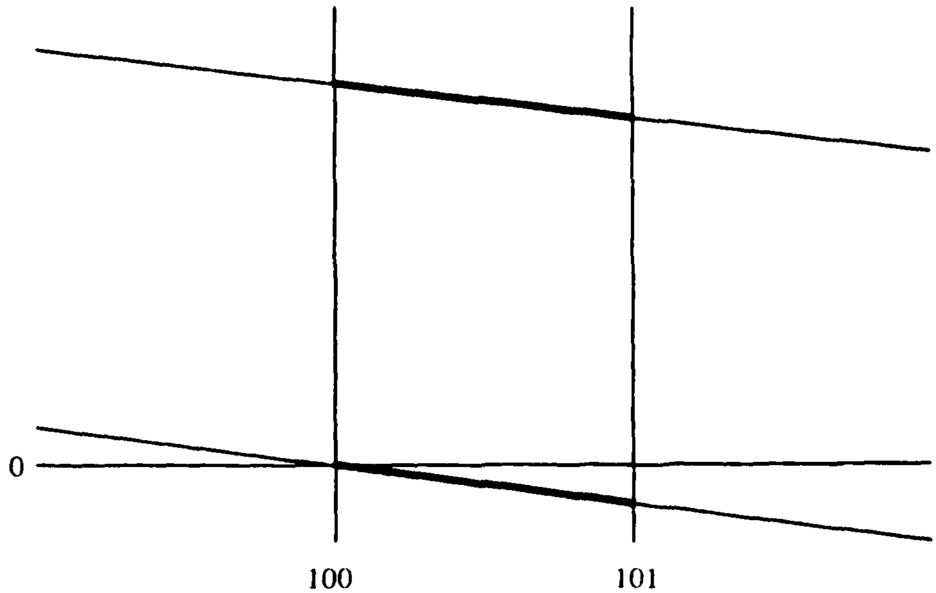


Figure 8.1 The power coefficients of these two very different linear polynomials differ by only 0.1%.

8. Condition

The **condition** of a basis measures how closely relative changes in the coefficients are matched by the resulting relative changes in the element represented. The closer the match, the better conditioned the basis is said to be. For example, the power basis $1, t, t^2, \ldots$ is not a good way to represent polynomials if we are interested in a positive interval $[a, b]$ with a/b close to 1. If, e.g., $[a, b] = [100, 101]$, then a 0.1% change in the power coefficients of the straight line $p : t \to t - 100$ can change its behavior on $[100, 101]$ by 100%; see Fig. 8.1.

If we use the appropriately shifted power basis, e.g., write p in the form $p(t) = \alpha + \beta(t - 100)$, then a .1% change in the coefficients α, β of this form produces a .1% change in the polynomial on the interval $[100, 101]$. The appropriately shifted power basis is often much better conditioned than the power basis. In this section, we discuss briefly the condition of the B-spline basis.

This requires us to bound the spline in terms of its B-spline coefficients and the B-spline coefficients in terms of the spline. The first turns out to be easy, while the second requires some work. Precisely, we are looking for constants $m > 0$ and M for which the inequalities

$$m \max_i |a_i| \le \max_t \left| \sum_i B_{ik}(t) a_i \right| \le M \max_i |a_i| \tag{8.1}$$

hold regardless of what the coefficient vector $a = (a_i)$ might be. Since the B-splines are *nonnegative and sum to 1* at any point, we have

$$\left| \sum_i B_{ik}(t) a_i \right| \le \sum_i B_{ik}(t) |a_i| \le \sum_i B_{ik}(t) \max_i |a_i| = \max_i |a_i|,$$

hence the second inequality always holds with $M = 1$. For the first inequality, we have to work a little harder.

Set $s := \sum_i B_{ik} a_i$. We know from Theorem 7 that

$$a_i = \lambda_{ik} s = \sum_{\nu=1}^{k} \frac{(-D)^{\nu-1} \psi_{ik}(\tau)}{(k-1)!} D^{k-\nu} s(\tau) \tag{8.2}$$

with τ some point which we can freely choose in the interval $[t_i, t_{i+k}]$. We now bound this sum in terms of $\max_t |s(t)|$.

Suppose that $\tau \in [t_l, t_{l+1}] \subset [t_i, t_{i+k}]$. Then, for some const_k depending only on k, and for all $p \in \pi_{<k}$ and all j,

$$|D^j p(\tau)| \le \mathrm{const}_k \, (\Delta t_l)^{-j} \max_{t_l \le t \le t_{l+1}} |p(t)| . \tag{8.3}$$

The existence of such a const_k follows for the case $\Delta t_l = 1$ from the fact that $\pi_{<k}$ is finite-dimensional, and from this it follows for arbitrary Δt_l by scaling. Since s agrees with some polynomial of degree $< k$ on $[t_l, t_{l+1}]$, we conclude that

$$|D^j s(\tau)| \le \mathrm{const}_k \, (\Delta t_l)^{-j} \max_{t_l \le t \le t_{l+1}} |s(t)| . \tag{8.4}$$

On the other hand, $\psi_{ik} = (t_{i+1} - \cdot)\dots(t_{i+k-1} - \cdot)$ is also a polynomial of degree $< k$, and

$$\max_{t_l \le t \le t_{l+1}} |\psi_{ik}(t)| \le \text{const}'_k |\Delta t_{l^*}|^{k-1} \tag{8.5}$$

for some const'_k which depends only on k and with $[t_{l^*}, t_{l^*+1}]$ a largest interval of that form in $[t_i, t_{i+k}]$. Therefore we choose $l = l^*$ and then obtain, from (8.3) with $p = \psi_{ik}$ and from (8.4), the bound

$$|D^{\nu-1}\psi_{ik}(\tau) D_{k-\nu} s(\tau)| \le (\text{const}_k)^2 \text{const}'_k \max_{t_i \le t \le t_{i+k}} |s(t)|.$$

Now sum these bounds over ν and divide by $(k-1)!$ to obtain

$$|a_i| = |\lambda_{ik} s| \le \text{const} \max_{t_i \le t \le t_{i+k}} |s(t)|,$$

with const depending only on k.

We have proved the following

Theorem 8. There exists a constant D_k depending only on k so that, for all knot sequences $\mathbf{t}$ and all $s \in S_{k,\mathbf{t}}$, and for all i,

$$|\lambda_{ik} s| \le D_k \max_{t_i \le t \le t_{i+k}} |s(t)|. \tag{8.6}$$

The best value for D_k is not known exactly but there is strong numerical evidence that $D_k \sim 2^{k-1}$. If we only consider *cardinal* splines, i.e., only uniform knot sequences, then the best value for D_k is known to be less than $(\pi/2)^k$.

Corollary. The inequalities (8.1) hold with $m = 1/D_k$ and $M = 1$.

9. Evaluation

In this section, we discuss the use of the recurrence relations (1.4) for the evaluation of a spline

$$s = \sum_i B_{ik} a_i \tag{9.1}$$

from its B-spline coefficients (a_i).

We already observed in (4.2) that the recurrence relations imply

$$s = \sum_i B_{ik} a_i = \sum_i B_{i,k-1} a_i^{[1]}. \tag{9.2}$$

with

$$a_i^{[1]} := (1 - \omega_{ik}) a_{i-1} + \omega_{ik} a_i. \tag{9.3}$$

Note that $a_i^{[1]}$ is not a constant, but is the straight line through the points (t_i, a_{i-1}) and (t_{i+k-1}, a_i). In particular, *$a_i^{[1]}(t)$ is a convex combination of a_{i-1} and a_i if $t_i \le t \le t_{i+k-1}$.*

After $k-1$-fold iteration of this procedure, we arrive at the formula

$$s = \sum_i B_{i1} a_i^{[k-1]},$$

which shows that

$$s = a_i^{[k-1]} \quad \text{on} \quad [t_i, t_{i+1}].$$

Algorithm 9. From given constant polynomials $a_i^{[0]} := a_i$, $i = j-k+1, \ldots, j$, (which determine $s := \sum_i B_{ik} a_i$ on $[t_j, t_{j+1}]$), generate polynomials $a_i^{[r]}, r = 1, \ldots, k-1$, by the recurrence

$$a_i^{[r+1]} := (1 - \omega_{i,k-r}) a_{i-1}^{[r]} + \omega_{i,k-r} a_i^{[r]}, \quad j-k+r+1 < i \le j. \tag{9.4}$$

Then $s = a_j^{[k-1]}$ on $[t_j, t_{j+1}]$. Moreover, for $t_j \le t \le t_{j+1}$, the weight $\omega_{i,k-r}(t)$ in (9.4) lies between 0 and 1. Hence the computation of $s(t) = a_j^{[k-1]}(t)$ via (9.4) consists of the repeated formation of convex combinations.

In the *cardinal* case (see Sec. 2, esp. (2.2-4)), the algorithm simplifies, as follows. Now

$$s =: \sum_i N_k(\cdot - i) a_i = \sum_i N_{k-1}(\cdot - i) a_i^{[1]} / (k-1).$$

with

$$a_i^{[1]} := (i + k - 1 - \cdot) a_{i-1} + (\cdot - i) a_i.$$

Hence

$$s = a_j^{[k-1]} / (k-1)! \quad \text{on} \quad [j, j+1],$$

with (9.4)$\mathbb{Z}$

$$a_i^{[r]} := (i + k - r - \cdot) a_{i-1}^{[r-1]} + (\cdot - i) a_i^{[r-1]}, \quad j - k + r < i \le j.$$

In the *Bernstein-Bézier* case (see Sec. 2, esp. (2.5-9)), all the nontrivial weight functions $\omega_{i,k-r}$ are the same, i.e.,

$$\omega_{i,k-r}(t) = t.$$

Thus, for

$$s = \sum_{\mu+\nu=h} B_{(\mu,\nu)} a_{(\mu,\nu)},$$

we get

$$s = a_{(0,0)} \text{ on } [0,1],$$

with (9.4)$_{\mathbb{B}}$

$$a_{(\mu,\nu)}(t) = (1-t)a_{(\mu+1,\nu)} + t a_{(\mu,\nu+1)}, \quad \mu+\nu = r;\ r = h-1,\ldots,0.$$

This is **de Casteljau's Algorithm** for the evaluation of the B-form.

10. Differentiation

In this section, we derive a formula for the B-spline coefficients of the derivative of a spline in terms of the B-spline coefficients of the spline.

By Corollary 1 to Theorem 6, the derivative Ds of a spline $s \in S_{k,\mathbf{t}}$ is again a spline with the same knot sequence but of one order lower. This means that, by Theorem 7, we can compute its B-spline coefficients (a'_i) by the formula

$$a'_i = \lambda_{i,k-1}(Ds)$$

provided we use $\tau \in [t_i, t_{i+k-1}]$.

To relate a' to a, we express $\lambda_{i,k-1}D$ as a linear combination of the functionals λ_{ik}, making use of the fact that λ_{ik} *depends linearly on* ψ_{ik},– recall the definition

$$\lambda_{ik} : f \mapsto \sum_{\nu=1}^{k} \frac{(-D)^{\nu-1}\psi_{ik}(\tau)}{(k-1)!} D^{k-\nu} f(\tau), \qquad (4.5b)$$

and that

$$(t_{i+k-1} - t_i)\psi_{i,k-1} = \psi_{ik} - \psi_{i-1,k}. \qquad (10.1)$$

These facts imply that

$$\begin{aligned}(\lambda_{ik} - \lambda_{i-1,k})f(\tau) &= \sum_{\nu=1}^{k} \frac{(-D)^{\nu-1}(\psi_{ik} - \psi_{i-1,k})(\tau)}{(k-1)!} D^{k-\nu} f(\tau)\\ &= (t_{i+k-1} - t_i)\sum_{\nu=1}^{k-1} \frac{(-D)^{\nu-1}\psi_{i,k-1}(\tau)}{(k-1)!} D^{k-\nu} f(\tau),\end{aligned}$$

the last equality by (10.1) and since $D^{k-1}\psi_{i,k-1} = 0$. On the other hand, directly from the definition (4.5b),

$$\begin{aligned}\lambda_{i,k-1}Df(\tau) &= \sum_{\nu=1}^{k-1} \frac{(-D)^{\nu-1}\psi_{i,k-1}(\tau)}{(k-2)!} D^{k-1-\nu} Df(\tau)\\ &= (k-1)\sum_{\nu=1}^{k-1} \frac{(-D)^{\nu-1}\psi_{i,k-1}(\tau)}{(k-1)!} D^{k-\nu} f(\tau).\end{aligned}$$

Comparison of these two displays shows that

$$\lambda_{i,k-1}D = \frac{k-1}{t_{i+k-1}-t_i}\left(\lambda_{ik} - \lambda_{i-1,k}\right). \tag{10.2}$$

Assuming that $B_{i,k-1} \neq 0$, i.e., that $t_i < t_{i+k-1}$, we can choose $\tau \in (t_i, t_{i+k-1}) = (t_{i-1}, t_{i+k-1}) \cap (t_i, t_{i+k})$. This yields

Algorithm 10. Compute the coefficients for $\sum a_i' B_{i,k-1} := D \sum a_i B_{ik}$ by

$$a_i' = \frac{a_i - a_{i-1}}{(t_{i+k-1}-t_i)/(k-1)}, \text{ if } t_i < t_{i+k-1}. \tag{10.3}$$

Remark What happens when $t_i = t_{i+k-1}$? In this case, $B_{i,k-1} = 0$, hence there is no need to calculate a_i'. To be precise, in this case, the spline $s = \sum_i B_{ik} a_i$ may not even be continuous at t_i, therefore $(Ds)(t_i)$ makes no sense. On the other hand, the left and the right limit, $(Ds)(t_i-)$ and $(Ds)(t_i+)$, always make sense, and the algorithm would provide all the a_j''s needed for their calculation. By applying the algorithm to the particular coefficient sequence $a = (\delta_{ij})$, we obtain the formula

$$DB_{ik} = \frac{k-1}{t_{i+k-1}-t_i} B_{i,k-1} - \frac{k-1}{t_{i+k}-t_{i+1}} B_{i+1,k-1}. \tag{10.4}$$

In terms of the alternative notations (1.9) for B-splines, this reads

$$DN_{ik} = M_{i,k-1} - M_{i+1,k-1}.$$

Since $\sum_i N_{ik} = 1$, this implies that

$$\int M_{i,k-1} = 1 \tag{10.5}$$

and so indicates why the particular normalization

$$M_{ik} := \frac{k}{t_{i+k}-t_i} B_{ik}$$

is of interest.

In the *cardinal* case, (10.3) reduces to

$$a_i' = a_i - a_{i-1} =: \nabla a_i \tag{10.3$_{\mathbb{Z}}$}$$

and (10.4) reads

$$DN_k = N_{k-1} - N_{k-1}(\cdot - 1). \tag{10.4$_{\mathbb{Z}}$}$$

On integrating this formula, we obtain

$$N_k(t) = \int_{t-1}^{t} N_{k-1}(\tau)d\tau \tag{10.6}$$

since both sides of (10.6) vanish for negative t. In terms of the convolution product

$$(f * g)(t) := \int f(t-\tau)g(\tau)d\tau$$

of two functions f and g, this gives the important formula

$$N_k = N_{k-1} * N_1(\cdot + 1). \tag{10.7}$$

This shows that $N_k(\cdot + k - 1)$ is the k-fold convolution product of N_1, i.e.,

$$N_k(\cdot + k - 1) = \underbrace{N_1 * N_1 * \ldots * N_1}_{k \text{ terms}}.$$

In the *Bernstein-Bézier* case, we get

$$D \sum_{\mu+\nu=h} B_{(\mu,\nu)} a_{(\mu,\nu)} = \sum_{\mu+\nu=h-1} B_{(\mu,\nu)} a_{(\mu,\nu)}$$

with $\qquad\qquad (10.3)_{\mathbb{B}}$

$$a_{(\mu,\nu)} = (\mu + \nu + 1)\big(a_{(\mu+1,\nu)} - a_{(\mu,\nu+1)}\big).$$

11. Knot insertion

In this section, we discuss the most important CAGD contribution to (univariate) spline theory, viz., the idea of knot insertion (a.k.a. subdivision). Since the spline order, k, will not change in this section, we will usually suppress it and write B_i instead of B_{ik}, ψ_i instead of ψ_{ik}, etc.

Simply put, knot insertion involves rewriting a given spline as a spline with a refined knot sequence, as can always be done by Corollary 2 of Theorem 6. Such a calculation is worthwhile since the B-spline coefficients are nearly equal to values of the spline at known points, and this is more nearly so when the knots are closer together. Here is the precise statement.

Theorem 11. If the spline $s = \sum_i B_i a_i$ is continuously differentiable, then

$$|a_i - s(t_i^*)| \leq \text{const } |\mathbf{t}|^2 \sup_t |D^2 s(t)| . \tag{11.1}$$

with

$$t_i^* := (t_{i+1} + t_{i+2} + \ldots + t_{i+k-1})/(k-1) \tag{4.7}$$

and

$$|\mathbf{t}| := \sup_i (t_{i+1} - t_i).$$

Proof Recall from Sec. 8 that

$$|a_i| = |\lambda_i s| \leq \text{ const } \max_{t_i \leq t \leq t_{i+k}} |s(t)|. \tag{8.6}$$

Further, recall from Sec. 4 (esp. (4.8)) that

$$\lambda_i p = p(t_i^*) \quad \forall p \in \pi_1.$$

Thus, choosing, in particular, $p := s(t_i^*) + (\cdot - t_i^*)Ds(t_i^*)$, the linear Taylor polynomial for s at t_i^*, we get

$$|a_i - s(t_i^*)| = |a_i - p(t_i^*)| = |\lambda_i(s-p)| \leq \text{ const } \max_{t_i \leq t \leq t_{i+k}} |(s-p)(t)|$$
$$\leq \text{ const } \frac{(t_{i+k} - t_i)^2}{8} \max_{t_i \leq t \leq t_{i+k}} |D^2 s(t)|. \quad \Box$$

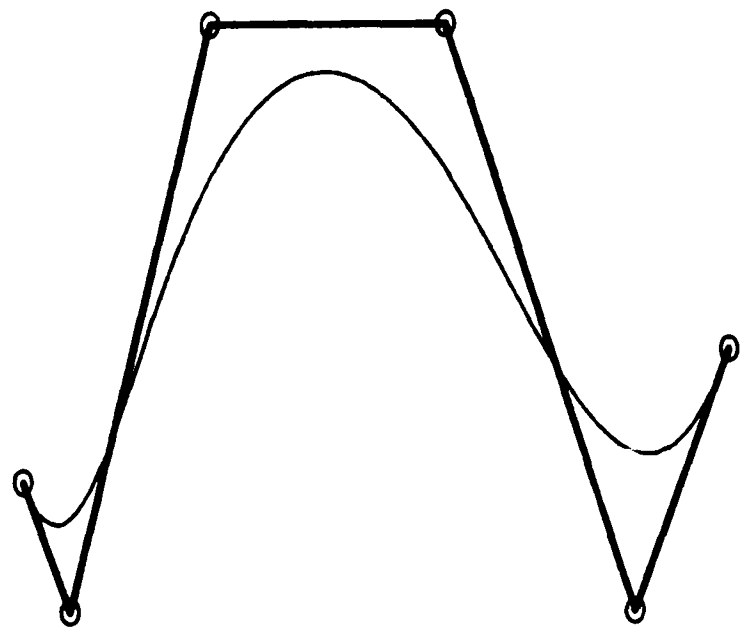


Figure 11.1 A cubic spline and its control polygon. The end knots are quadruple.

This suggests consideration of the **control polygon** associated with the representation $\sum_i B_i a_i$ of the spline s as an element of $S_{\mathbf{t}}$. This control polygon will be denoted by

$$C_{a,\mathbf{t}}.$$

It is the broken line or piecewise linear function with vertices $P_i := (t_i^*, a_i)$. For, the theorem implies that the control polygon will be close to s if $|\mathbf{t}|$ is small. Here is the precise statement.

Corollary. Let $C_{a,\mathbf{t}}$ be the control polygon associated with the representation $\sum_i B_i a_i$ of the continuous spline s as an element of $S_{\mathbf{t}}$. Then

$$\sup_t |s(t) - C_{a,\mathbf{t}}(t)| \le \text{ const } |\mathbf{t}|^2 \sup_t |D^2 s(t)|. \tag{11.2}$$

Proof Let $t_i^* \le t \le t_{i+1}^*$ and let p be the linear polynomial which agrees with s at t_i^* and t_{i+1}^*. Then

$$|s(t) - p(t)| \le |t_{i+1}^* - t_i^*|^2/8 \max_{t_i^* \le \tau \le t_{i+1}^*} |D^2 s(\tau)|,$$

while

$$|p(t) - C_{a,\mathbf{t}}(t)| \le \max\{|s(t_i^*) - a_i|, |s(t_{i+1}^*) - a_{i+1}|\} \le \text{ const } |\mathbf{t}|^2 \max_\tau |D^2 s(\tau)|$$

by the theorem. |||

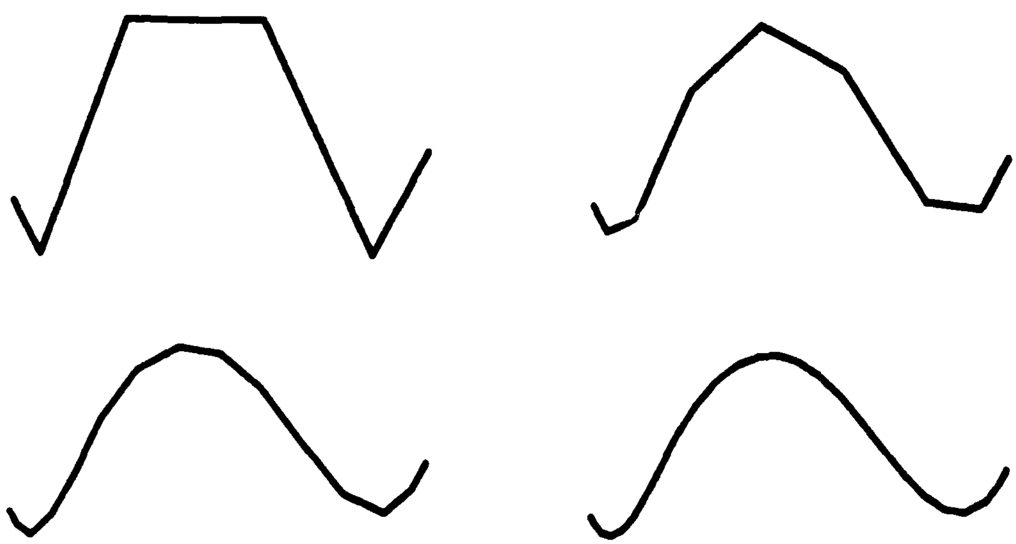


Figure 11.2 The control polygon of Fig. 11.1 and three midpoint refinements.

This shows that the control polygon $C_{a,\mathbf{t}}$ converges to the spline s as we refine the knot sequence $\mathbf{t}$. Since the typical graphical equipment only draws broken lines, anyway, this makes it attractive to construct refined control polygons for a spline.

For this, we need to know how to compute, from its B-spline coefficients a_i as an element of $S_{\mathbf{t}}$, the B-spline coefficients $\hat{a}_i$ for the spline s with respect to a refined knot sequence $\hat{\mathbf{t}}$. By Theorem 7, this is a question of comparing the corresponding $\hat{\lambda}_i$ with λ_i. Since the dual functional

$$\lambda_i : f \mapsto \sum_{\nu=1}^{k} \frac{(-D)^{\nu-1}\psi_i(\tau)}{(k-1)!} D^{k-\nu} f(\tau) \tag{4.5b}$$

depends linearly on ψ_i, this requires nothing more than to express

$$\hat{\psi}_i = (\hat{t}_{i+1} - \cdot) \cdots (\hat{t}_{i+k-1} - \cdot)$$

as a linear combination of the ψ_i.

This is particularly easy when $\hat{\mathbf{t}}$ is obtained from $\mathbf{t}$ by adding just one knot, say the point $\hat{t}$. Then

$$\hat{\psi}_i = \begin{cases} \psi_i, & t_{i+k-1} \le \hat{t}; \\ \psi_{i-1}, & \hat{t} \le t_i, \end{cases}$$

hence there is some actual computing necessary only for $t_i < \hat{t} < t_{i+k-1}$. For this case.

$$\begin{aligned} \alpha\psi_{i-1} + \beta\psi_i &= (t_{i+1} - \cdot) \cdots (t_{i+k-2} - \cdot)\left[\alpha(t_i - \cdot) + \beta(t_{i+k-1} - \cdot)\right] \\ &= \hat{\psi}_i \end{aligned}$$

provided $\alpha(t_i - \cdot) + \beta(t_{i+k-1} - \cdot) = (\hat{t} - \cdot)$, i.e.,

$$\alpha = 1 - \omega_i(\hat{t}) \text{ and } \beta = \omega_i(\hat{t}).$$

Since $\hat{t}_i = t_i < \hat{t} < t_{i+k-1} = \hat{t}_{i+k}$, we can choose τ in the definition (4.5b) in the interval $(\hat{t}_i, \hat{t}_{i+k}) = (t_{i-1}, t_{i+k-1}) \cap (t_i, t_{i+k})$. This proves

Algorithm 11. If the knot sequence $\hat{\mathbf{t}}$ is obtained from the knot sequence $\mathbf{t}$ by addition of the point $\hat{t}$. then the coefficients $\hat{a}_i$ for the spline s with respect to the refined knot sequence are given by

$$\hat{a}_i = \begin{cases} a_i, & \text{if } t_{i+k-1} \le \hat{t}; \\ (1 - \omega_i(\hat{t}))a_{i-1} + \omega_i(\hat{t})a_i, & \text{if } t_i < \hat{t} < t_{i+k-1}; \\ a_{i-1}, & \text{if } \hat{t} \le t_i. \end{cases} \tag{11.3}$$

Observe that $\omega_i(\hat{t}) \in [0.1]$ when $t_i < \hat{t} < t_{i+k-1}$. and thus the coefficients $\hat{a}$ are convex combinations of the coefficients a.

This algorithm has the following very pretty graphical interpretation.

Corollary. The refined control polygon $C_{\hat{a},\hat{t}}$ can be thought of as having been obtained by interpolation at its vertices to the original control polygon $C_{a,t}$. i.e.,

$$C_{\hat{a},\hat{t}}(\hat{t}_i^*) = C_{a,t}(\hat{t}_i^*) \text{ for all } i. \tag{11.4}$$

Proof Consider the straight line $p : t \mapsto t$. It is a spline and, by (4.8),

$$p = \sum_i B_i t_i^*,$$

i.e., (t_i^*) is its B-spline coefficient sequence with respect to the knot sequence $\mathbf{t}$. In particular, it is its own control polygon, i.e., $C_{\mathbf{t}^*,\mathbf{t}} = p$, regardless of what the knot sequence $\mathbf{t}$ might be. This implies that (11.3) also holds with every a replaced by t^*. |||

This says that the point $\hat{P}_j := (\hat{t}_j, \hat{a}_j)$ lies on the segment $[P_{j-1}, P_j]$ and cuts this segment in the ratio $(\hat{t} - t_j) : (t_{j+k-1} - \hat{t})$. This is illustrated in Figure 11.3 for the control polygon of Figure 11.1.

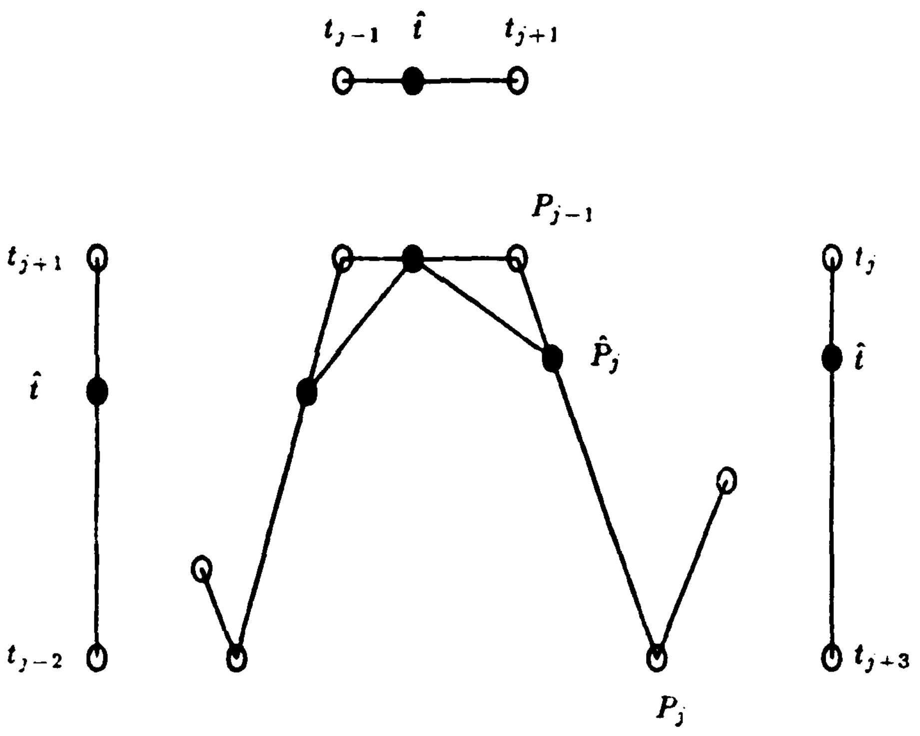


Figure 11.3 Insertion of $\hat{t} = 2$ into the knot sequence
$\mathbf{t} = (0,0,0,0,1,3,5,5,5,5)$, with $k = 4$.

If $r := \#\hat{t} \le k - 1$, then, after just $(k - 1 - r)$-fold insertion of $\hat{t}$, we obtain a knot sequence $\hat{\mathbf{t}}$ in which the number $\hat{t}$ occurs exactly $k - 1$ times. This means that there is exactly one B-spline for that knot sequence which is not zero at $\hat{t}$. Hence it must equal 1 at $\hat{t}$ and its coefficient must provide the value of s at $\hat{t}$. This makes it less surprising that the calculations in Algorithms 9 and 11 are identical.

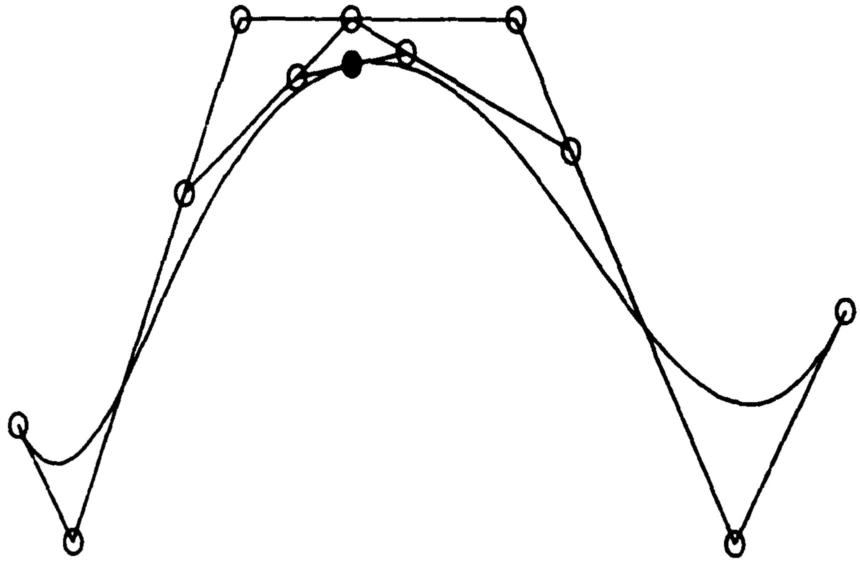


Figure 11.4 The cubic spline and its control polygon from Figure 11.1 and the sequence of control polygons generated by three-fold insertion of the same knot. (The finest control polygon differs from its predecessor only by an additional vertex point.)

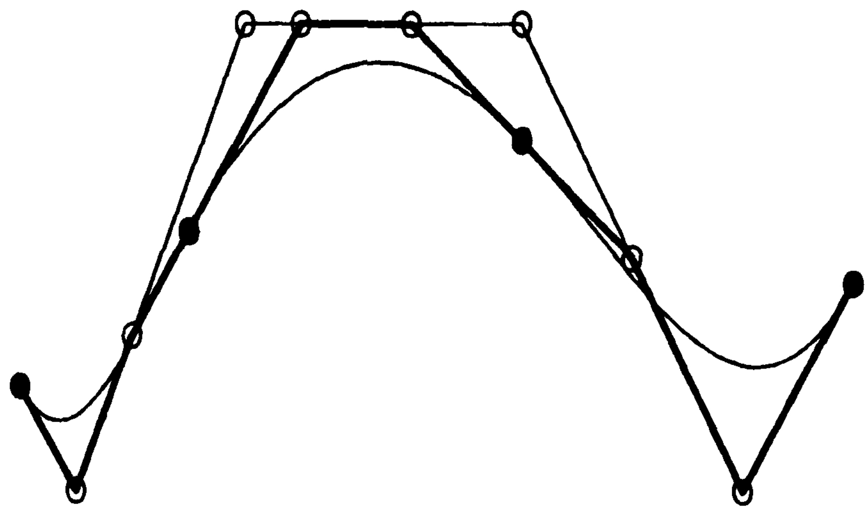


Figure 11.5 Conversion to B-net by $(k-2)$-fold insertion of each knot

Conversion to B-net Let $\mathbf{t}'$ be the refined knot sequence which contains each of the knots in $\mathbf{t}$ exactly $k-1$ times. Then each corresponding B-spline B'_{jk} is nonzero on just one knot interval, hence coincides there with a properly shifted and scaled element of the **Bernstein basis**. The k B-spline coefficients a'_j associated in this way with a knot interval therefore provide the coefficients in the B-form for the polynomial with which the spline agrees on that knot interval. The coefficient sequence (a'_i), or the control polygon $C_{a',t'}$, are called the **B-net** for the given spline. It can be obtained by inserting each knot t_i of the spline $k - 1 - \#t_i$ times. The process can be speeded up slightly by inserting first every other knot, and, in a second round, inserting the remaining knots. The latter insertion process is then entirely local and depends only on the ratio of the two knot intervals containing the knot being inserted.

While the formulas do simplify for the *cardinal* case, they are not of much use in that form since insertion of one knot into the sequence $\mathbf{t} = \mathbb{Z}$ would destroy the uniformity of the knot sequence. But it makes good sense to develop formulas for inserting the same number of uniformly spaced knots into every interval $[i, i+1]$ since this produces again a *uniform* knot sequence. Because of its practical importance, we treat this case separately, in the next section.

12. Knot insertion for cardinal splines

In this section, we consider knot refinement for cardinal splines, i.e., splines with a uniform knot sequence. Here it is desirable to have the refined knot sequence again uniform. We restrict attention to the case that the given knot sequence is $\mathbf{t} = \mathbb{Z}$. This is no real restriction since an arbitrary uniform knot sequence can always be written in the form $\alpha + \beta\mathbb{Z}$ for appropriate scalars α and β, and if s is a spline with that knot sequence, then $s(\alpha + \beta\cdot)$ is a spline with the knot sequence $\mathbb{Z}$.

If we insert $m - 1$ uniformly spaced knots into every knot interval of $\mathbb{Z}$, then the refined knot sequence is $\hat{\mathbf{t}} = \mathbb{Z}/m$. The corresponding B-splines $\hat{B}_i$ are

$$\hat{B}_i = \hat{N}_k(\cdot - i),$$

with

$$\hat{N}_k(t) := N_k(mt)$$

an appropriately scaled version of the standard cardinal B-spline N_k. This makes it trivial to determine $\hat{a}_i$ in case $k = 1$. Since

$$N_1 = \hat{N}_1 + \hat{N}_1(\cdot - 1) + \ldots + \hat{N}_1(\cdot - m + 1), \tag{12.1}$$

we find for this case that

$$\hat{a}_{im+j} = a_i \quad \text{for } j = 0, \ldots, m - 1.$$

The formula for general order k is obtained from this with the aid of the convolution formula

$$N_k = N_{k-1} * N_1(\cdot - 1) \tag{10.7}$$

from Sec. 10, as follows. We define

$$s_r := \sum_i N_r(\cdot - i) a_i =: \sum_i \hat{N}_r(\cdot - i) a_{ir}, \quad r = 1, \ldots, k. \tag{12.2}$$

Then $\hat{a}_i = a_{ik}$, and, from (10.7) and (12.1),

$$\begin{aligned} s_{r+1} = s_r * N_1(\cdot + 1) &= \Big(\sum_i \hat{N}_r(\cdot - i)a_{ir}\Big) * \sum_{j=1}^m \hat{N}_1(\cdot - j) \\ &= \sum_i \sum_{j=1}^m \underbrace{\hat{N}_r(\cdot - i) * \hat{N}_1(\cdot + j)}_{\hat{N}_{r+1}(\cdot - i - j + 1)/m} \; a_{ir} \\ &= \sum_i \sum_{j=1}^m \hat{N}_{r+1}(\cdot - i)/m \;\; a_{i+j-1,r}. \end{aligned}$$

Here, we have used the following consequence of the convolution formula (10.7):

$$\begin{aligned} \hat{N}_{r-1}(\cdot - \alpha) * \hat{N}_1(\cdot - \beta) &= \int N_{r-1}(m(\cdot - \tau) - \alpha)N_1(m\tau - \beta)d\tau \\ &= \int N_{r-1}(m\,\cdot - \sigma + \beta - \alpha)N_1(\sigma)d\sigma/m \\ &= \hat{N}_r(\cdot + \beta - \alpha). \end{aligned}$$

We conclude that

$$a_{i,r+1} := (a_{i,r} + a_{i+1,r} + \ldots + a_{i+m-1,r})/m, \quad \text{for } r > 0. \tag{12.2}$$

Here is the full algorithm.

Algorithm 12. Given the B-spline coefficients $a = (a_i)$ of $s \in S_{k,\mathbb{Z}}$, its B-spline coefficients $\hat{a} = (\hat{a}_i)$ with respect to the refined knot sequence $\mathbb{Z}/m$ can be computed as follows:

$$a_{mi+j,1} := a_i, \quad j = 0, \ldots, m-1;$$

$$a_{i,r} = \sum_{j=0}^{m-1} a_{i+j,r-1}/m, \quad r = 2, \ldots, k;$$

$$\hat{a}_i := a_{ik}.$$

In practice, one would use the algorithm repeatedly with $m = 2$ rather than once with a larger m. For, the computational cost is

$$nm(k-1)((m-1)A + D),$$

with n the number of coefficients to start with, and A and D the cost of one addition, respectively division. If, e.g., the targeted refinement is to have $2^\mu n$ coefficients, then the cost ratio of the choice $m = 2^\mu$ versus the use of μ applications of the algorithm, each time with $m = 2$, is

$$\frac{2^\mu((2^\mu - 1)A + D)}{(2 + 2^2 + \ldots + 2^\mu)(A + D)} \sim 2^{\mu-1}A + D/2.$$

In addition, even though the repeated application, with $m = 2$, takes roughly twice as many divisions, these are just divisions by 2.

13. Shape preservation

In this section, we use knot insertion to prove the shape preserving property of B-splines. Roughly speaking, this property says that a spline has the same shape as its control polygon.

We begin with the

convex hull property If $t_j \le t < t_{j+1}$, then $s(t)$ is a convex combination of the k B-spline coefficients $a_{j-k+1}, \ldots, a_j$.

which follows from Algorithm 9 or directly from the facts that B-splines are nonnegative (Sec. 1) and add up to 1 at every point (see (4.6)).

For a statement of the full shape preserving property, we recall that

$$S^-(a)$$

is the standard notation for the number of (strong) sign changes in a sequence a. Thus

$$S^-(1,-1,1,-1) = 3, \quad S^-(1,0,1,-1) = 1, \quad S^-(0,0,0,0) = 0.$$

Theorem 13. Variation diminution $S^-(s) \le S^-(a)$: i.e., with $x_1 < \cdots < x_r$ arbitrary,

$$S^-(s(x_1), \ldots, s(x_r)) \le S^-(a).$$

Proof. Recall from Sec. 11 that $s(x_1), \ldots, s(x_r)$ is a *subsequence* of the sequence $\bar{a}$ of coefficients for s with respect to the refined knot sequence $\bar{t}$ which contains each x_i at least $k-1$ times. Hence it is sufficient to prove that $S^-(\bar{a}) \le S^-(a)$. But this follows once we know that $S^-(\hat{a}) \le S^-(a)$, with $\hat{a}$ obtained by (11.3), i.e., by insertion of just one knot. For this simple case, though, the conclusion is immediate if we think of the construction of $\hat{a}$ from a as occurring in two steps: In the first step, we insert $\hat{a}_i$ between a_{i-1} and a_i, and this does not increase the number of sign changes since each $\hat{a}_i$ is a *convex* combination of its neighbors a_{i-1} and a_i in that new sequence. In the second step, we pull out $\hat{a}$ as a subsequence, and this may only lower the number of sign changes.

Corollary. Shape preservation A spline crosses any straight line no more often than does its control polygon. In particular, if the control polygon is monotone (convex), then so is the spline.

Proof Let s be the spline and p the straight line. Then $S^-(s-p)$ is the number of times the spline crosses the straight line. Since $s-p$ is a spline, this is bounded by $S^-(a-b)$, with a, b the B-spline coefficients of s, resp. p with respect to $\mathbf{t}$, and this equals the number of times the control polygon $C_{a,\mathbf{t}}$ crosses the control polygon for p. But, as we observed in Sec. 11, the control polygon for the straight line p is p itself. This proves the general statement.

For the particulars, recall that a (continuous) function is monotone if and only if it crosses any any constant function at most once, and that a function is convex if it crosses any straight line at most twice (dipping first below and then rising above the line in case it crosses it twice). ▯

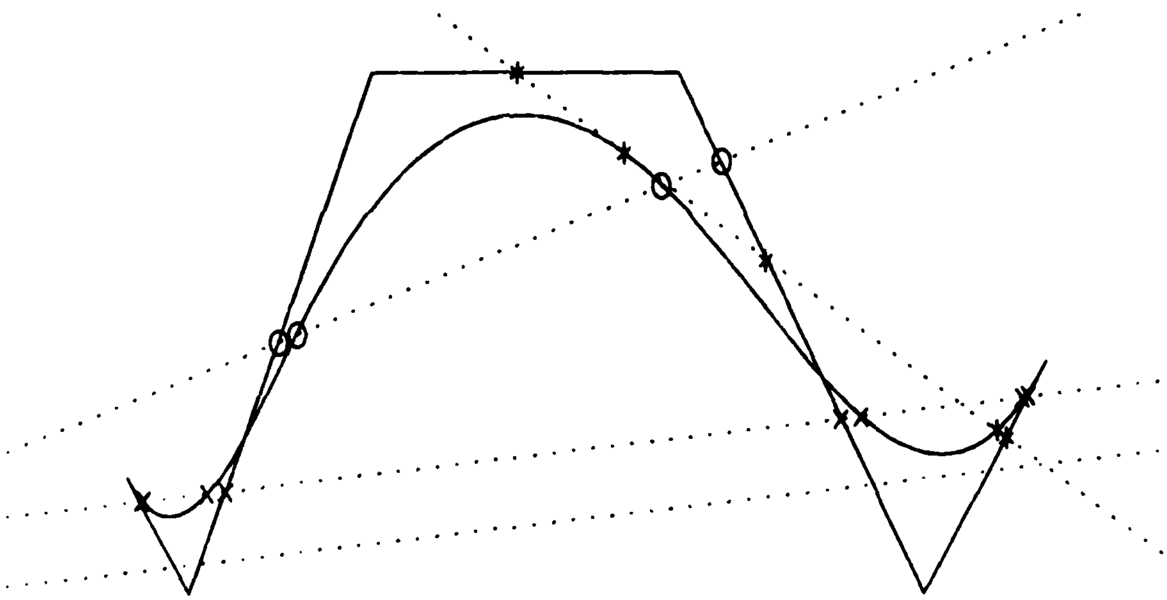


Figure 13.1 A cubic spline, its control polygon, and various straight lines intersecting them. The control polygon *exaggerates* the shape of the spline. The spline crossings are bracketed by the control polygon crossings.

REFERENCES

[BBB85] R. H. Bartels. J. C. Beatty and B. A. Barsky (1985), An Introduction to the Use of Splines in Computer Graphics, SIGGRAPH'85, San Francisco.

[BFK84] W. Böhm, G. Farin and J. Kahmann (1984), A survey of curve and surface methods in CAGD. *Computer Aided Geometric Design* **1**. 1-60.

[B76] C. de Boor (1976). Splines as linear combinations of B-splines, in *Approximation Theory II*, G. G. Lorentz, C. K. Chui and L. L. Schumaker eds., 1-47.

[B78] C. de Boor (1978). *A Practical Guide to Splines*, Springer-Verlag. New York.

[BH86] C. de Boor and K. Höllig (1986). B-splines without divided differences, MRC TSR #2896: in *Geometric Modeling*, G. Farin ed., SIAM, xxx-xxx.

[Scho69] I. J. Schoenberg (1969). Cardinal interpolation and spline functions, *J. Approximation Theory* **2**. 167-206.

[Schu81] L. L. Schumaker (1980), *Spline Functions*. J. Wiley, New York.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2952	2. GOVT ACCESSION NO. AD-A179 773	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) B(asic)-SPLINE BASICS		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Carl de Boor		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53705		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 3 - Numerical Analysis and Scientific Computing
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE August 1986
		13. NUMBER OF PAGES 34
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

splines
B-splines

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An expository and illustrated treatment of the basic B-spline theory as derived from the recurrence relations.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END
11-86
DTIC